Calibration and Image Reconstruction Techniques for FMCW MIMO Radars with Consideration of Antenna Radiation Patterns

Kalibrations- und Bildrekonstruktionsverfahren für
FMCW-MIMO-Radare unter Berücksichtigung der
Antennenabstrahlungscharakteristik

Daniel Gotzens

Masterarbeit vorgelegt dem

Lehrstuhl für Radarsystemtechnik
Institut für Hochfrequenztechnik
RWTH Aachen University
December 6, 2023

In Kooperation mit

indurad GmbH Aachen

Abstract

Increased image quality can be achieved by implementing backprojection on a millimeter-wave TDM FMCW MIMO Radar. In this thesis, a backprojection-based imaging algorithm is proposed and implemented for a radar sensor. This radar sensor's functionality is tested, the sensor is calibrated, and then the algorithm is compared to the current FFT-based algorithm as well as MUSIC. It is found that near-field accuracy of the image generated with backprojection and MUSIC is increased, and that real-time capability can only be guaranteed for backprojection and the current FFT-based approach.

Contents

Chapter 1

Introduction

1.1 Theoretical Background

1.1.1 Signal Model

Single Channel FMCW

A single channel consists of a transmit antenna and a receive antenna. The transmit antenna sends a so-called chirp of duration T_{chirp}, which is a sinusoid with linearly increasing frequency. The signal $x_{TX}(t)$ send by the transmit antenna is reflected by an ideal point scatterer at position $\vec{r}_S$ and then received at the receive antenna as $x_{RX}(t)$. The propagation delay τ can be calculated using the speed of light c_0, and the locations of the receive and transmit antennas $\vec{r}_{RX}$ and $\vec{r}_{TX}$:

$$\tau = \frac{\|\vec{r}_{TX} - \vec{r}_S\| + \|\vec{r}_{RX} - \vec{r}_S\|}{c_0} \tag{1.1}$$

Using a complex representation for the in-phase and quadrature components of the signal, the transmit and receive signal can be formulated for $t \in [0, T_{chirp}]$:

$$x_{TX}(t) = A_0 e^{j(\omega_0 t + \frac{1}{2}\dot{\omega}t^2 + \phi_0)} \tag{1.2}$$

$$x_{RX}(t) = A(\vec{r}_S)x_{TX}(t - \tau) \tag{1.3}$$

$$\tag{1.4}$$

The received signal is then mixed with a copy of the transmitted signal (1.2) and a low-pass filter is applied. The resulting signal $y(t)$ is called *intermittent*

frequency signal.

$$y(t) = \mathrm{LP}\left\{ x_{RX}(t) \cdot x_{TX}(t) \right\} \tag{1.5}$$

$$= \mathrm{LP}\left\{ A_0 e^{j(\omega_0 t + \frac{1}{2}\dot{\omega} t^2)} \cdot A(\vec{r}_S) A_0 e^{j(\omega_0(t - \tau(\vec{r})) + \frac{1}{2}\dot{\omega}(t - \tau(\vec{r}))^2)} \right\} \tag{1.6}$$

$$= A_0^2 A(\vec{r}_S) e^{j(\frac{1}{2}\dot{\omega}\tau^2(\vec{r}_S) - \omega_0\tau)} \cdot \mathrm{LP}\left\{ e^{j(2\omega_0 t + \frac{1}{2}\dot{\omega} t^2 - \dot{\omega}\tau t)} \right\} \tag{1.7}$$

$$\approx G(\vec{r}_S) e^{-j\dot{\omega}\tau t} \tag{1.8}$$

The fact that the IF-signal contains all the information – i.e. the IF signal's frequency directly corresponds to the target's distance – explains the main advantage of this technology. The carrier frequency can be orders of magnitude higher than the intermittent frequency, which drastically reduces the requirements for the subsequent signal processing, while retaining the improved resolution due to the smaller wavelenghts of the carrier frequency.

FIND QUOTE:"GHz resolution for MHz processing"

To locate a target in the cross-range dimensions, a single-channel FMCW-radar can be used to scan in multiple directions, by either rotating the antennas, redirecting their beam with rotating mirrors, or with beamforming antenna arrays. In any case, this requires highly directive antennas and also increases size, weight and cost of a radar sensor.

MIMO FMCW

Multiple-input multiple-output radar benefits from increased diversity and signal power. If N_{TX} transmit antennas and N_{TX} receive antennas are employed, $K = N_{TX} \cdot N_{RX}$ channels are available. To differentiate the signals from each other, a multiplexing technique has to be chosen. Options include time division multiplex, frequency division multiplex and code division multiplex.

In TDM, multiple access is achieved by the transmit antennas all send one after another, while all receive antennas receive simultaneously. In FDM, simultaneous transmission is made possible by subdividing the bandwidth and assigning a different frequency range to each antenna. That means that TDM allows for higher bandwidths for each transmission, while FDM allows higher transmission durations.

In CDM, both simultaneous transmission and use of the entire bandwidth is made possible by using a different waveform to each channel. However, processing at the carrier frequency is required to differentiate the signals from another, as opposed to TDM and FMD, where all processing can be done at the intermittent frequency range.

Depending on the application, a compromise has to be found between the advantages and drawbacks of each method. There are also methods available that combine aspects of these three basic paradigms, such as OFDM and Hadamard-Coding.[citation needed]

3

Once the received signals are demultiplexed, the ideal receive signal for antenna pair $k \in \{0, 1, ...K-1\}$:

$$y_k(t) = G_k(\vec{r}_S)e^{-j\dot{\omega}\tau_k(\vec{r}_S)t} \tag{1.9}$$

Note that both the gain and the propagation delay may differ from channel to channel.

In reality, the scene can consist of multiple and expansive scatterers, that reflect the transmitted signals at different intensities. which is summarized as a locational reflectivity $F_k(\vec{r})$. [1] Also, interference and electric noise may be present in each channel, which we summarize as $n_k(t)$. Thus, the overall IF-signal is:

$$y_k(t) = \iiint F_k(\vec{r})G_k(\vec{r})e^{-j\dot{\omega}\tau_k(\vec{r})t} \, d\vec{r} + n_k(t) \tag{1.10}$$

$$\tag{1.11}$$

After sampling the signal at sampling intervals T_s such the sampling frequency $f_s = \frac{1}{T_s}$ is sufficiently high: $2f_s > \frac{1}{2\pi}(\omega_0 + \dot{\omega}T_{chirp})$, and with M samples such that $MT_s < T_{chirp}$, the sampled IF-signal can be defined as:

$$y_k[m] = y_k(t = mT_s), \text{for } m \in \{0, 1, ..M-1\} \tag{1.12}$$

1.1.2   Calibration

1.1.3   Image Reconstruction

Image reconstruction is an inverse problem where the locational reflectivity of the scene $F(\vec{r})$ has to be estimated from the received signals $y_k[m]$. Multiple approaches are available; in the following, three will be presented.

Discrete Fourier Transform

The discrete fourier transform can be implemented with high efficiency, and many CPUs even include silicone-based implementations [citation needed]. In this approach, the DFT is applied over three dimensions of the input signal, obtaining a discrete output signal in spherical coordinates whose amplitude is an estimate of the locational reflectivity.

For each input channel, the range of a target can be estimated by applying the DFT over time. The resulting spectrum's peak corresponds to the target:

$$\mathcal{F}_m\{y_k[m]\}(\Omega) = \sum_{m=0}^{M-1} e^{-j2\pi\frac{m\Omega}{M}} y_k[m] \tag{1.13}$$

$$= G_k(\vec{r}_S)\delta(\Omega - \dot{\omega}\tau_k(\vec{r}_S)T_s) \tag{1.14}$$

---

[1]The index k is introduced here to take obstructed visibility into account: from the point of view of one channel, two scatterers may be visible simultaneously, while from the point of view of another, one might obstruct the other's visiblity.

In order to understand how information on the direction of a target can be extracted from the channel data, we consider an ideal $1 \times K$ horizontal uniform linear array (ULA) where the spacing is exactly $d = \frac{\lambda_0}{2}$, with $\lambda_0 = \frac{c_0}{f_0}$. The antennas are located at $\vec{r}_{TX} = \vec{0}$ and $\vec{r}_{RX,k} = (kd, 0, 0)^T$. A scatterer located at $\vec{r}_S = (r_S \sin\theta_S, r_S \cos\theta_S, 0)^T$ reflects the transmitted radar waves with an intensity of A_S. Then, their runtime across the array is:

$$\tau_k = \frac{1}{c_0} \left(\|\vec{r}_{TX} - \vec{r}_S\| + \|\vec{r}_{RX} - \vec{r}_S\| \right) \tag{1.15}$$

$$\tag{1.16}$$

In far-field conditions, the target is far enough ($r \gg Kd$) away for the reflected wavefronts to be planar. That means that the runtime can then be approximated as such:

$$\tau_k \approx 2r_S + kd \sin\theta_S \tag{1.17}$$

In equation 1.7 it can be seen that the locational gain $G(\vec{r})$ contains a phase shift depending on the runtime of the waves:

$$G_k(\vec{r}) = A_S A_k(\vec{r}) e^{j(\frac{1}{2}\dot{\omega}\tau_k^2 - \omega_0 \tau_k)} \tag{1.18}$$

$$\approx A_S A_k(\vec{r}) e^{-j\omega_0 \tau_k} \tag{1.19}$$

Assuming the attenuation along the path is channel-independent, the signals can all be considered copies of each other:

$$y_k(t) = A_S A_k(\vec{r}) e^{-j\omega_0 \tau_k} e^{-j\dot{\omega}\tau_k t} \tag{1.20}$$

$$= A_S A_k(\vec{r}) e^{-j\frac{\omega_0}{c_0}(2r_S + kd\sin\theta_S)} e^{-j\frac{\dot{\omega}}{c_0}(2r_S + \overbrace{kd\sin\theta_S}^{\ll 2r_S})t} \tag{1.21}$$

$$= A_S A_0(\vec{r}) e^{-j\frac{\omega_0}{c_0}2r_S} e^{-j\frac{\dot{\omega}}{c_0}2r_S t} e^{-j\frac{\omega_0}{c_0}kd\sin\theta_S} \tag{1.22}$$

$$= y_0(t) e^{-j\frac{\omega_0}{c_0}kd\sin\theta_S} \tag{1.23}$$

$$\Rightarrow y_k[m] = y_k(t = mT_S) = y_0[m] e^{-j\frac{\omega_0}{c_0}kd\sin\theta_S} \tag{1.24}$$

Applying the DFT accross the ULA yields:

$$\mathcal{F}_k\{y_k[m]\}(\Omega) = \mathcal{F}_k\{y_0[m] e^{-j\frac{\omega_0}{c_0}kd\sin\theta_S}\}(\Omega) \tag{1.25}$$

$$= y_0[m] \cdot \delta\left(\Omega - \frac{\omega_0}{c_0}(d\sin\theta_S)\right) \tag{1.26}$$

$$= y_0[m] \cdot \delta\left(\Omega - \pi \sin\theta_S\right) \tag{1.27}$$

$$\tag{1.28}$$

The azimuth angle θ_S can be extracted from the signal supplied by a horizontal ULA. Analogously, the elevation angle ϕ_S can be obtained with a vertical

ULA. If a $1 \times K$ array is used, where the K receive antennas form a uniformly spaced grid, successive DFTs across the rows and columns of this grid yield two dimensions. However, the same can be achieved with fewer antennas in a MIMO configuration.

The *virtual array* of a MIMO array is a corresponding SIMO array. It has the same number of channels: if the original array is $N_{TX} \times N_{RX}$, the virtual array has $1 \times K$ channels, where $K = N_{TX} \cdot N_{RX}$.

The virtual transmit and receive antenna associated with channel $k = 0$ are placed in the origin of the virtual array's coordinate system. The other receive antennas are then placed such that the displacement between them and the transmit antenna is the same as it was between the original MIMO array's corresponds transmit and receive antenna. It is possible for virtual antennas to be at the same location.

Overall, a 3D image in range, azimuth, and elevation is generated by calculating the DFT over time, and the DFTs over the rows and columns of the virtual array. For this to work, the scatterer needs to be distant enough for the wavefronts to be planar, and the virtual array's grid needs to be uniformly spaced with $d = \lambda_0 / 2$ spacing.

Backprojection

Compared to the DFT-based approach, backprojection takes fewer approximations and requirements on the array to work, while using a similar amount of computation. The approach works by correlating the input signal $y_k[m]$ to the theoretical signal $s_k[m, \vec{r}]$ of an ideal scatterer at different locations. The mean correlation of all channels to the theoretical signal is then used as an estimate for the locational reflectivity:

$$\hat{F}(\vec{r}) = \frac{1}{K} \sum_{k=0}^{K-1} s_k[m, \vec{r}] \star y_k[m] \tag{1.29}$$

$$= \frac{1}{K} \sum_{k=0}^{K-1} \sum_{m=0}^{M} s_k^*[m, \vec{r}] y_k[m] \tag{1.30}$$

Using the signal model from (1.9) yields:

$$\hat{F}(\vec{r}) = \frac{1}{K} \sum_{k=0}^{K-1} \sum_{m=0}^{M} G_k^*(\vec{r}) e^{+j\dot{\omega}\tau_k(\vec{r}_S)mT_s} y_k[m] \tag{1.31}$$

To reduce the computational intensity of this algorithm, calculating the inner sum (over m) can be rewritten as an inverse discrete fourier transform (IDFT):

$$\hat{F}(\vec{r}) = \frac{1}{K} \sum_{k=0}^{K-1} G_k^*(\vec{r}) \sum_{m=0}^{M} e^{+j\dot{\omega}\tau_k(\vec{r}_S)mT_s} y_k[m] \tag{1.32}$$

$$= \frac{1}{K} \sum_{k=0}^{K-1} G_k^*(\vec{r}) \sum_{m=0}^{M} e^{j\Omega m} y_k[m] \Big|_{\Omega = \dot{\omega}\tau_k(\vec{r}_S)T_s} \tag{1.33}$$

$$= \frac{1}{K} \sum_{k=0}^{K-1} G_k^*(\vec{r}) \mathcal{F}_m^{-1}\{y_k[m]\} (\Omega = \dot{\omega}\tau_k(\vec{r}_S)T_s) \tag{1.34}$$

MUSIC

The Multiple Signal Classification (MUSIC) algorithm can also be used to estimate the locational reflectivity of a scene. It operates on the time-domain fourier transform of the IF-signal, and makes similar far-field approximations as the DFT-based approach. The abstract signal model for MUSIC is:

$$\mathbf{y}(t) = \mathbf{A} \cdot \mathbf{s}(t) + \mathbf{n}(t) \tag{1.35}$$

Here, $\mathbf{y}, \mathbf{n} \in \mathbb{C}^K$, $\mathbf{A} \in \mathbb{C}^{K \times Z}$, and $\mathbf{s} \in \mathbb{C}^Z$. Z is the number of voxels in the output image and K the number of receive channels. For example, if the output image consist of $X \times Y \times Z$ cuboid voxels, then $Z = X \cdot Y \cdot Z$.

Thus, the support matrix $\mathbf{A}$ is a linear transform from the locational reflectivity $\mathbf{s}$ to the expected input signal vector $\mathbf{y}$. Each collumn vector $\mathbf{a}_z$ of the support matrix $\mathbf{A}$ therefor corresponds to the expected input signal vector caused by a point source.

The MUSIC algorithm revolves around the correlation matrix of its input signal $\mathbf{R_{yy}}$. Assuming the a stationary scene with zero-mean noise of covariance $\mathbf{C_{nn}}$, it follows that

$$\mathbf{R_{yy}} = \mathrm{E}\{\mathbf{y}\mathbf{y}^H\} \tag{1.36}$$

$$= \mathbf{A}\mathbf{R_{ss}}\mathbf{A}^H + \mathbf{C_{nn}}, \text{ with } \mathbf{R_{ss}} := \mathrm{E}\{\mathbf{s}\mathbf{s}^H\} \tag{1.37}$$

Assume that $\mathbf{R_{ss}}$ is nonsingular with rank q and that $\mathbf{A}$ has full rank. If $\mathbf{R_{yy}}$ has p eigenvalues, then the smallest $p - q$ of them are all σ^2, and their corresponding eigenvectors – i.e. the collumns of $\mathbf{C_{nn}}$ – are all orthogonal to the support vectors $\mathbf{a}_z$.

This property is key to the MUSIC algorithm. The metric used to generate an image is the projection of $\mathbf{a}_z$ onto the $\mathbf{C_{nn}}$. Due to their orthogonality, the projection of support vectors corresponding to a signal source will be zero. The image intensity at voxel z is thus defined computed as the normalized inverse square magnitude of this projection:

$$P_{MUSIC}[z] = \frac{\mathbf{a}_z^H \mathbf{a}_z}{\mathbf{a}_z^H \hat{C}_{nn}^H \hat{C}_{nn} \mathbf{a}_z} \tag{1.38}$$

The input signals are often highly correlated, due to phenomena such as multipath propagation or inter-channel crosstalk. This unfortunately means that nonsingularity of $\mathbf{R_{ss}}$ cannot always be guaranteed. A preprocessing step is required to "decorrelate" the signals and thereby making $\mathbf{R_{ss}}$ singular again.

While early schemes, such as the "3/4in plywood" spacial dither algorithm by Widrow *et al.* [CITE] consisted of mechanically moving the receive antenna array orthogonal to the look direction, preprocessing can also be done after receiving the signal.

Spacial smoothing, as proposed by [CITE], improves the correlation matrix's eigenstructure by

1.2   Physical Setup

Chapter 2

Calibration

In order to achieve high fidelity imaging, any systematic offsets of the radar sensor must be compensated. In this chapter, the systematic errors of the radar are observed and classified (sec. 2.1). With knowledge of their stability and overall characteristics, an online calibration method is implemented and its performance analised (sec. 2.2).

2.1 Stability Analysis

The systematic errors described in section ¡¿ manifest by distorting amplitude and phase of each channel's deramped signal. Static calibration is feasible if the systematic errors are not affected by changes of the following variables: system restart, time elapsed since system restart, system temperature and the deramped signal's frequency.

2.1.1 Setup

The sensor is placed in a low-reflection environment and a corner reflector is placed at boresight in front of the sensor at a distance of roughly $1.60\,\mathrm{m}$ (cf. 2.1). The time data collected in all channels is recorded every minute. The temperature readings of the sensor's CPU, FPGA and radar frontend are also recorded every minute (cf. 2.2). The experiment is repeated with a different ramp slope $\dot{\omega}$ in order to achieve a different frequency of the deramped signal, since

$$f_{IF} = \dot{\omega} \cdot \frac{2r_{refl}}{c} \tag{2.1}$$

Due to the geometry of the setup, the runtime of each transmitted wavefront should be identical. Thus, the ideal deramped signal should be of a single, constant frequency and without inter-channel phase differences. In practice, the reflector peak will have a certain bandwidth, and other peaks at higher and
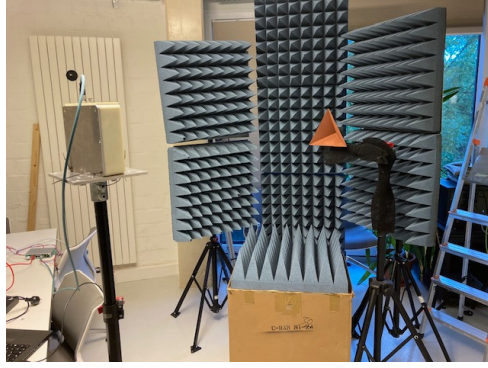
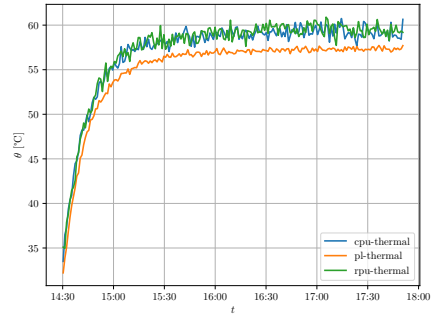Figure 2.1: Measurement Setup

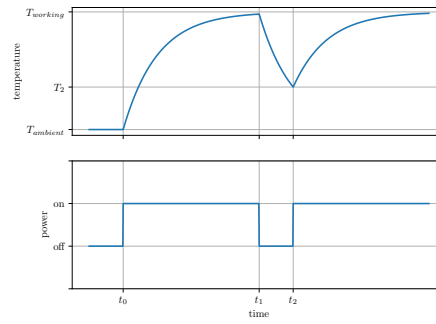Figure 2.2: Measured System Temperature After Startup

Figure 2.3: Expected System Temperature over Time for Interrupted Power

lower frequencies will be present due to incomplete shielding and/or unwanted reflections. Also, the reflector peak may wander if the setup geometry moves. For analysis, system runtime and temperature cannot be considered independent variables, as illustrated in figure 2.3: The system starts at ambient temperature, heating up and approaching a stable operating temperature on turning on (t_0), and cooling back down after turning off (t_1).
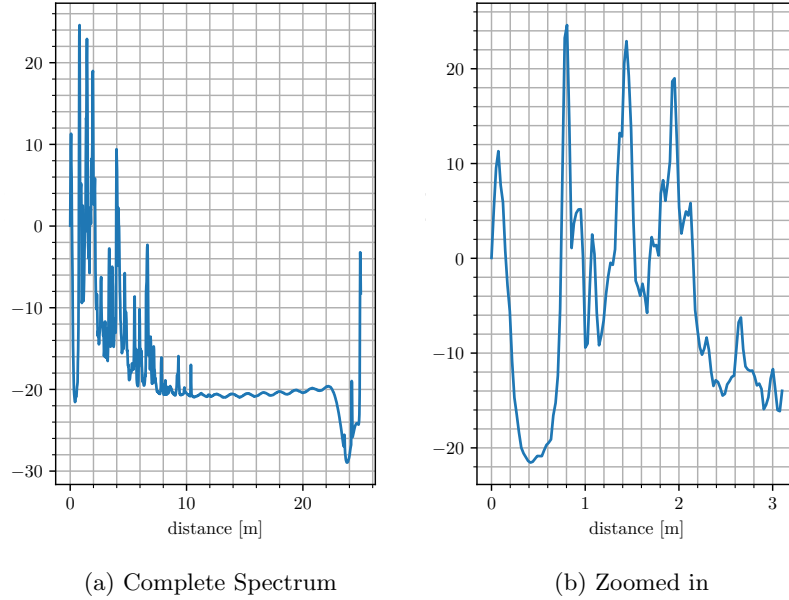
2.1.2   Results

(a) Complete Spectrum   (b) Zoomed in

Figure 2.4: Mean Intensity Spectrum

Preprocessing and Analysis

Because of the aforementioned imperfections in the experiment, some preprocessing is required to analize the systematic offsets present in the radar signal. Multiple additional peaks in the spectrum are visible in figure 2.4; indeed, the maximum peak is not even caused by the reflector. It is therefore necessary to limit the analysis to only the FFT-bin at the maximum of the peak caused by the reflector.

The measurement environment also exhibits some minor changes in temperature and humidity, which can result in the geometry to shift by a few millimeters. This shows by the spectral maximum shifting over time, as seen in figure 2.5.
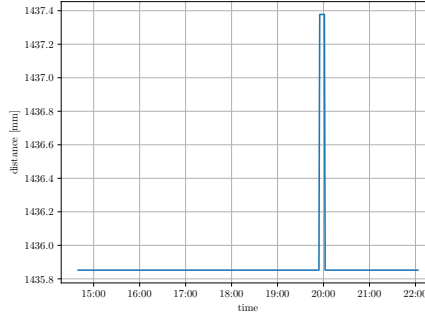
11

Figure 2.5: Change in reflector distance as measured by iMCR

In the following analysis, the main metric is the complex ratio of the maximum FFT-bin to its initial value. Using a logarithmic representation of this ratio, it can be represented as a level difference [dB] and a phase difference [°].

Effects of System Temperature and Runtime

Multiple measurements have indicated that, while the amplitude rarely varies by more than $1\,\text{dB}$, the phase is not as stable over time. Typically, the mean phase drifts by up to $50°$ in the hours after system startup, with the rate of change reducing after around four hours. However, it has to be noted that there is no clear correlation between phase drift and temperature: the phase continues to drift after the system temperature stabilizes; the reduction in drift only occurs hours after the system has reached a stable temperature of approximately $60\,°\text{C}$.

Effects of Self-Calibration

As described in the specification, the AWR2243P-chips undergo a self-calibration upon initialization. This initialization can be triggered by either restarting the entire system or by re-writing the configuration registers on the radar chips. Indeed, this initial calibration can be observed in the data (cf. 2.7). After the connection with the radar has been re-established, the following effects are visible:

- slightly increased level: the level in each channel increases by approximately $0.3\,\text{dB}$

- increased incoherence: after re-starting the system, the drift in both level and phase is distributed more broadly

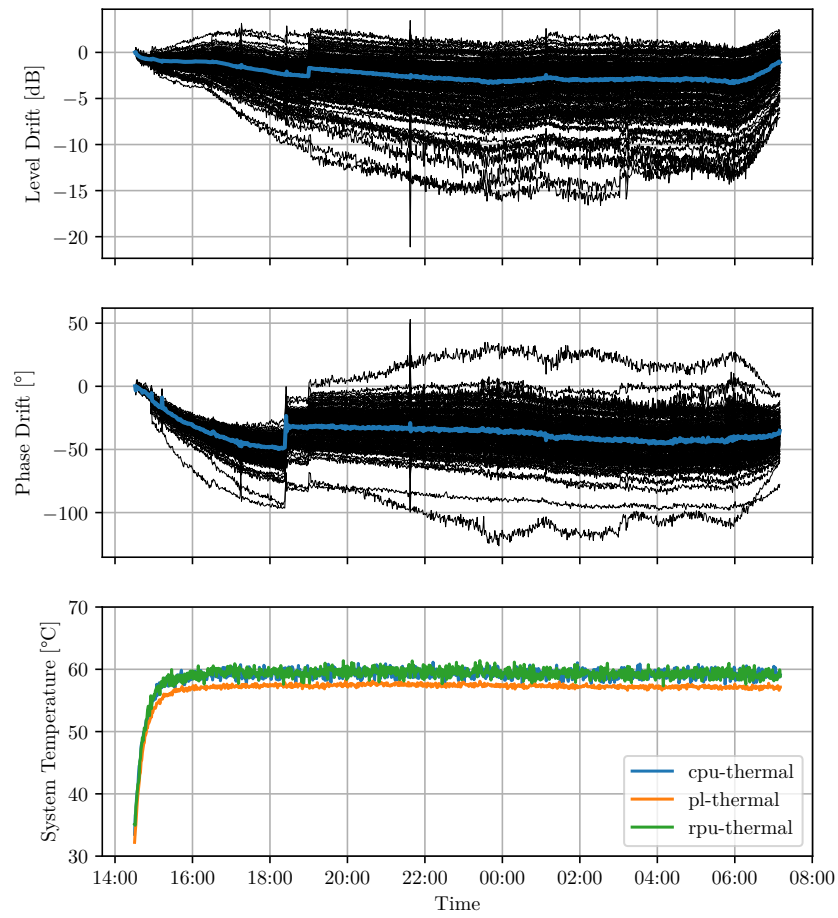- mean phase: the mean phase drift changes by up to $5°$
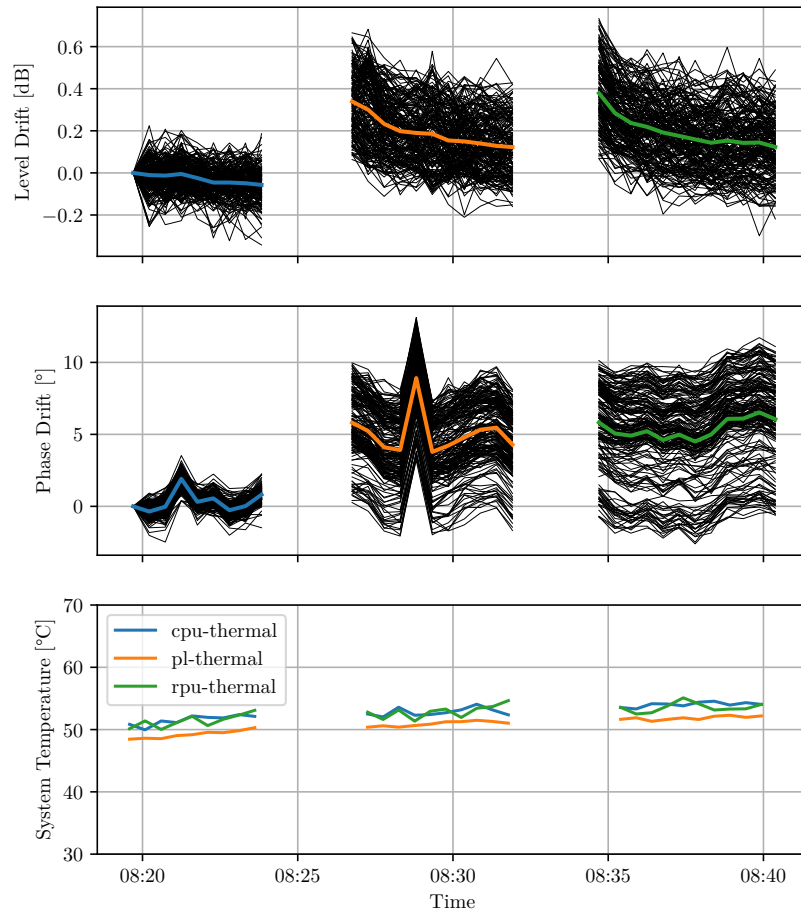
Figure 2.6: Recorded drift over night

Figure 2.7: Recorded drift and temperature with system restart

Differences between Antenna Pairs

Comparing the amplitude recorded by all antenna pairs, no channel in particular stands out. When looking at the different channels phase drifts, it can be seen that the array becomes less coherent over time. The distribution of phase shift across channels seems roughly gaussian, with increasing variance over time. No antenna pair seems more prone to incoherence than the rest: across measurements, the maximum outlier (in terms of divergence from the mean phase) cannot be associated with a chip or antenna.

Effects of the Recorded Frequency and Angle

- ?

Conclusion

-

2.2   Calibration

Chapter 3

Image Reconstruction

3.1 Implementation

3.2 Results

Chapter 4

Results

List of Figures